OFFICE OF ECONOMICS WORKING PAPER
U.S. INTERNATIONAL TRADE COMMISSION

TRADE FRICTIONS AND WELFARE IN THE GRAVITY MODEL: HOW MUCH OF THE ICEBERG MELTS?

Edward J. Balistreri
Russell H. Hillberry
Office of Economics
U.S. International Trade Commission

November 2001

Trade Frictions and Welfare in the Gravity Model:
How Much of the Iceberg Melts?

by
Edward J. Balistreri
U.S. International Trade Commission

Russell H. Hillberry
U.S. International Trade Commission and
Purdue University

November 4, 2001

Abstract: Theoretical gravity models predict the bilateral trade pattern quite well. Does this justify their use in more ambitious exercises, such as estimating the welfare losses attributable to trade barriers? No. The models' weaknesses are not obvious in reduced form econometric work, but become quite apparent when we explore the implied general equilibrium. We simulate a model of U.S.-Canada interregional trade based on econometric estimates from the literature. We find that the estimates imply that over 45% of U.S. and Canadian output is consumed in transit. Removing the border cost does not change this result significantly, as the model requires implausibly large iceberg transportation costs to fit observable data. We conclude that the theoretical underpinnings of the gravity equation are far from settled, and welfare analyses based on the estimated equations cannot be validated.

Contact information: Edward J. Balistreri, Office of Economics, U.S International Trade Commission, 500 E Street S.W., Washington, DC 20003. Phone: (202) 205-3265 e-mail: ebalistreri@usitc.gov

I. Introduction

The gravity model has long been considered a robust empirical relationship without firm theoretical foundations. The development of differentiated products models with gravity-like predictions has led some authors to conclude otherwise, and even to suggest that such models should replace traditional methods of welfare and policy analysis. Careful consideration of model implications suggests important weaknesses in the value of these models as analytical tools. The models' quantitative predictions are sharply at odds with available evidence on the resource costs of transportation and the geographic dispersion of prices. Not surprisingly, given these failures, the models' welfare predictions are highly implausible and easily discarded.

Structural gravity models are appealing for two reasons: 1) empirical estimation produces a well fitting regression and 2) the models imply a parsimonious interpretation of robust econometric estimates. For example, in one of the best-known econometric applications of the gravity model, McCallum (1995) found that trade among Canadian provinces exceeded gravity-adjusted province-state trade by a factor of twenty. Structural gravity models suggest that the McCallum estimate implies that the costs of trading across the U.S. Canadian border are equivalent to *ad valorem* tariffs of between 75 and 300%.[1]

In the most formal econometric application of the theory so far, Anderson and van Wincoop (2001) (henceforth A-vW) estimate a structural version of Anderson's (1979) theoretical gravity model with inter-regional trade and commodity flow data from Canada and the U.S. A-vW are confident in the value of their technique, for it helps reduce the size of the estimated border effect. But a careful reading of A-vW hints at the underlying weakness of this and other structural gravity models - they rely too heavily on consumer behavior to motivate the trade pattern. It is only by imputing implausible geographic differences in consumer price indices that A-vW can explain bilateral trade flows. To wit, A-

---

[1] These estimates apply the theoretical derivation outlined below to the McCallum estimate of the border effect. The tariff equivalent estimate depends upon the choice of a taste parameter, the elasticity of substitution among regions' aggregate production bundles. The figures here rely on our assumption that that the elasticity of substitution lies somewhere between 5 and 2, respectively. We justify this range below.

vW's "solution" to the border puzzle requires that retail prices in Canada exceed prices in the United

States by at least 33 percent, and perhaps as much as 318 percent.[2]

In this paper, we highlight additional shortcomings of structural gravity models through

numerical simulation of implied interregional trade patterns in the U.S. and Canada. We follow the

structure of Anderson, using econometric estimates from A-vW to evaluate quantitative predictions

implied by a strict model-based interpretation of U.S.-Canadian interregional trade flows. We rely on the

A-vW estimates because the authors make the boldest claims about the validity of the structural approach

and the value of model-based structural estimation. We remain uncomfortable with the assumptions

implicit in A-vW's reconciliation of U.S. Commodity Flow Survey data with interprovincial and

province-state trade data (see Hillberry (2001)). However, our critique of structural gravity models need

not rest on A-vW's results. The iceberg melt implied by most any gravity-based econometric estimates,

including McCallum's, is excessive.

Under a range of parameterizations, we find substantial evidence to discredit the model. Most

notable is the following: *transport costs must consume over 45 percent of U.S. and Canadian output if the*

Anderson model is to replicate the interregional trade pattern. This finding is so significant that it drives

most of the subsequent results. In this context, large welfare gains from removing the border are not

surprising, as border costs consume 10% of Canadian output. Given that total border and distance costs

consume a total of 54% of Canadian output, rescuing an additional 1/5 of the Canadian output available

for consumption produces enormous welfare benefits.

Following Anderson (1979) and most of the gravity literature, we model transport and border

costs with the Samuelsonian iceberg form; a portion of output melts in transit. While the most damning

results reported below are expressed in terms of iceberg melt, they cast doubt on the model's other

---

[2] These estimates also assume an elasticity of substitution between 5 and 2, respectively. The calculation takes the ratio of the "multilateral resistance terms" in row 1 of table 3 in A-vW, and raises it to the power $(1/1-\sigma)$, where σ is the elasticity of substitution.

assumptions as well. Structural gravity models assume that all output is tradable; that there is no intermediate goods trade; and that each good is produced in only one location. It is the combination of these assumptions that puts an excessive burden on iceberg melt as a determinant of bilateral trade patterns. In short, our results put the entire model in jeopardy, not just the iceberg assumption.

One might plausibly wonder if our critique depends on the restricted set of regions (Canadian provinces and a subset of U.S. states) used in A-vW's regression analysis. In fact, the restricted set of regions is artificially favorable to the structural model. In the Anderson framework, adding regions reduces each region's relative demand for its own output, raising its propensity to trade. Because larger trade volumes imply more iceberg melt (especially when the added regions are at a distance and across borders), adding regions would raise our estimates of iceberg melt. Our results appear to be a general critique of structural gravity models.

The structure of the paper is as follows. The following section includes a review of the literature on gravity and the border effect. The third section explains the model derivation, calibration, and includes a brief description of our data sources. Section IV reports results, contrasting them with observable data and common sense observation. Section V offers an assessment of the gravity literature and a guide to future research. Section VI concludes.

II. Literature review

The literature using the gravity model to estimate a border effect in North American interregional trade begins with McCallum (1995). In a simple log-linear gravity regression, McCallum finds that 1988 trade among Canadian provinces exceeded province-state trade by more than a factor of twenty. Helliwell (1996) and Anderson and Smith (1999a, 1999b) use updated versions of the Canadian data, finding slightly smaller, though still large, border effects. Using an alternative data source, the 1993 Commodity Flow Survey data, Hillberry (1998) finds a border effect of over 20 in U.S. commodity flows.

These authors focused largely on estimating magnitudes, without explicitly interpreting them in a structural framework.

In subsequent literature, authors have attempted to apply the insights of structural gravity models to the reduced form estimates of McCallum. Models by Anderson (1979), Bergstrand (1985), Deardorff (1998) and Krugman (1980) - which assume complete specialization, CES utility over differentiated products, and an *ad valorem*, iceberg form of transportation and border costs - all predict a gravity-like trade pattern. The models also suggest a rather straightforward interpretation of McCallum's border dummy coefficient, that it represents the product of an elasticity of substitution among differentiated products and an *ad valorem* border cost. Evans (1999) and Obstfeld and Rogoff (2000) argue explicitly for this interpretation of border effects. A-vW are firmly within this tradition.

One of the primary lessons of structural gravity models is that interregional trade costs can lead the relative price of the household consumption bundle to differ substantially over space, with central regions enjoying low average prices while remote regions suffer from the negative effects of high average prices. While price data at a regional level is generally not available, several authors (beginning with Wei (1996)) have included "remoteness" terms, which capture a region's GDP-weighted average distance from all regions in the sample, as proxies for regional price indices.[3] A-vW's innovation relative to this literature is to make explicit the non-linearities that a strict interpretation of the structural model implies about the construction of the proxy price term.[4]

A recent strand of the literature has offered a number of cogent critiques of structural gravity models as a means of understanding the data. Haveman and Hummels (1999) show that the complete

---

[3] Hummels (1999) removes the influence of price terms and other region specific idiosyncrasies from the regression by including origin and destination region-specific fixed effects. For most applications, this would appear to be the preferred econometric approach to identifying border-type barriers.

[4] It should be stressed that none of these papers, including A-vW, use actual price data; they all rely on proxies. The proxy terms are weighted averages of output and distance terms that appear elsewhere in the standard regression. It is worth noting that when studies of international trade explicitly include price data, as in Bergstrand (1985), the inclusion of price data does little to affect the estimated sign or magnitude of the coefficients on the gravity variables. Conclusions that rely on specifications with constructed price proxies should be accepted with a grain of salt.

specialization assumption critical to these models is not consistent with certain aspects of international data, and argue that gravity-like trade patterns can arise out of a model with incomplete specialization and homogenous goods. Evans (2001) argues that fixed costs are an important source of border effects. Hillberry (2001) shows that the border induces changes in the composition of trade, a fact not predicted by structural models and not easily reconciled with them. Hillberry and Hummels (2000) argue that intermediate goods trade is an important aspect of interregional trade, and that the co-location of firms trading intermediates can magnify very small border costs.

We are quite sympathetic to these critiques, and expect they will soon be more widely appreciated. Our purpose is to show that the structural models are an inappropriate way to interpret aggregate bilateral trade data. Rather than test each assumption individually, as the literature above does, we use the model to elicit quantitative predictions, and compare these predictions to observable data. It is the model's failure to credibly predict prices and transport costs, two variables that are critical to its use in welfare analysis, that discredits it.

III. Model derivation and calibration technique

The model follows Anderson (1979). For expositional purposes, we reproduce it here, with some notational adjustments to facilitate subsequent discussion of our calibration procedure. The set-up is straightforward. Consumers have preferences over goods differentiated by region of origin. Iceberg transportation costs induce substitution of consumption toward varieties produced nearby. Border costs induce substitution of consumption toward domestic varieties. The model is calibrated to regional output data and a matrix of bilateral distances, using parameter estimates from A-vW.

The general equilibrium framework

From utility maximization we set the consumer price indexes equal to the unit expenditure functions in each region:

$$CPI_j = \left(\sum_k \left(b_i FOB_i t_{ij} \right)^{(1-s)} \right)^{1/(1-s)} \tag{1}$$

CPI_i is the true cost of living index in region i. FOB_i is the price received in region i for units

of output. The distribution parameter (b_i) indicates homogeneous tastes across regions. The

trade cost factor is given by t_{ij}; the proportion of the shipment that melts is $(1 - 1/t_{ij})$.

The next set of conditions require market clearance for the country specific commodities,

and the composite commodities:

$$e_i^0 - \sum_j \left(\frac{Y_j}{FOB_i} \left[\frac{b_i t_{ij} FOB_i}{CPI_j} \right]^{1-s} \right) = 0 ; \tag{2}$$

$$a_i^0 U_i - \left(Y_i \Big/ CPI_i \right) = 0 . \tag{3}$$

Each region's initial (or benchmark) endowment of product is given by e_i^0. The scale parameter

a_i^0 equals benchmark income over the benchmark equilibrium price index, which is a convenient

normalization of utility. The final condition is that income for each region balances:

$$Y_i = FOB_i e_i^0 . \tag{4}$$

Together, conditions (1) through (4) are a complete multi-region general equilibrium that can be

solved numerically for relative prices, regional utility, and income levels.[5] As an artifact of the

---

standard regression. It is worth noting that when studies of international trade explicitly include price data, as in Bergstrand (1985), the inclusion of price data does little to affect the estimated sign or magnitude of the coefficients on the gravity variables. Conclusions that rely on specifications with constructed price proxies should be accepted with a grain of salt.

[5] For computation we represent the system as a mixed complementarity problem (Rutherford [1995]). With n regions, the system includes $4n$ equations (1, 2, 3, and 4) and $4n$ unknowns (Y_i, U_i, CPI_i, FOB_i). Only relative prices are determined, however, so one of the market clearance conditions is removed (by Walras' law) and we assign the associated price as the numeraire.

equilibrium, we can report the individual bilateral trade flow from region i to j (X_{ij}) as a function

of the equilibrium prices and the income level. These are given by the demand functions;

$$X_{ij} = \left(\frac{Y_j}{FOB_i} \left[\frac{b_i t_{ij} FOB_i}{CPI_j} \right]^{1-s} \right). \tag{5}$$

Calibration

To calibrate the system (1) through (4) we utilize A-vW's solution given identical

preferences and symmetric trade costs. Furthermore, we apply the econometrically estimated

coefficients from A-vW, and the income and distance data, to generate the benchmark

equilibrium consistent with fitted trade flows.

Following A-vW, the symmetric loglinear trade cost factor between i and j is given by:

$$t_{ij} = b_{ij} d_{ij}^r . \tag{6}$$

b_{ij} is the border cost factor, equal to one plus the tariff equivalent. d_{ij} is the bilateral distance

and r is the elasticity of transport costs with respect to distance. Estimation of the theoretical

gravity equation presented by A-vW provides the estimates $(1-s)r = -0.79$ and

$(1-s)\ln b = -1.65$, so the trade costs and calibration of the model are conditional upon a given

elasticity of substitution (s).

Another point that is generally ignored by the literature is the real implication of scaling

the distance measure. Changing the scale of distance has no relative effects, but it does affect the

implicit normalization. Hummels (1999) regresses freight charges on $r_0 d^{r_1}$, where ρ_0 is a scalar

that determines the level, as opposed to the shape, of distance costs. He finds that ρ_0 varies

across commodities, and is often significantly different than 1. In our framework, it can be

shown that gross trade flows and the border's effect on welfare are independent of how distance

is scaled. Decomposing the value of trade flows into net consumption and transport melt portions, however, depends on the given normalization.

Aware of this, we make an assumption that is most favorable to the model. We normalize the distance measures on the shortest measured pair, in this case New Jersey to New Jersey trade covers only 11 miles. We divide all of our absolute distance data by a scalar equal to 11. This ensures that for the shortest pair there is no melt; the trade cost factor is one for trade within New Jersey. This minimizes our estimate of the amount of income devoted to trade costs subject to the constraint that trade costs must be positive.[6]

While we present a range of substitution elasticities for completeness, we are not agnostic about the choice of this key parameter. To our knowledge, an elasticity of substitution among fully aggregated production bundles has not been estimated directly in the literature. However, it is a stylized fact that the elasticity of substitution falls with aggregation (Hummels [1999], Gallaway et al., [2001]). This suggests that Hummels' mean estimate of 4.79 in 1-digit SITC data should be an upper bound. A more convincing approach, in our view, is to apply Hummels' estimate of the distance elasticity of freight costs ($r = 0.27$) to A-vW's estimate of $(1-s)r = -0.79$. This suggests a σ of approximately 4. We settle on a range of 2 to 5 as the most plausible parameterizations of the model.[7]

To give an example of the trade cost factors implied by A-vW's estimates Table 1 presents these for a Canadian province and a US state for a range of substitution elasticities.[8] We can see from Table 1 that the cost factors are extremely large even with very large substitution rates. The indication is that in order for Alberta to get one unit of product from Alabama they need to purchase 299 to 4.16 units at the factory gate, given a substitution rate of 2 or 5

---

[6] When we do not scale distance, the transport sector consumes as much as 70% of output.
[7] It is plausible that other distance-related costs, such as communication and information costs, raise the true value of ρ above 0.27, thereby reducing our preferred estimate of σ even further. We appeal to the Hummels figure because it allows a direct comparison with the implied size of the transport sector.

respectively. These high markups are not fully realized in the price indices because regions can move their net consumption away from imports with high trade costs. Even the own cost factors (t_{ii}) seem quite high relative to actual transport markups. The cost factor on the Alberta-Alberta pair is 1.75 and for Alabama-Alabama it is 1.68 (for an elasticity of 5).

---

[8] We only present the trade cost factors for these two regions purely for brevity. All other cost factors are available on request.

Table 1

Trade Cost Factors for Alberta and Alabama
(t_{ij}; one indicates zero trade cost)

	Alberta			Alabama		
	$\sigma=2$	$\sigma=5$	$\sigma=10$	$\sigma=2$	$\sigma=5$	$\sigma=10$
Province						
Alberta	9	1.75	1.28	299	4.16	1.88
British Columbia	18	2.06	1.38	339	4.29	1.91
Manitoba	28	2.31	1.45	221	3.86	1.82
New Brunswick	67	2.86	1.59	239	3.93	1.84
Newfoundland	78	2.97	1.62	324	4.24	1.90
Nova Scotia	70	2.89	1.60	251	3.98	1.85
Ontario	54	2.71	1.56	158	3.54	1.75
Prince Edward Island	69	2.88	1.60	260	4.01	1.85
Quebec	59	2.77	1.57	198	3.75	1.80
Saskatchewan	18	2.05	1.38	253	3.99	1.85
State						
Alabama	299	4.16	1.88	8	1.68	1.26
Arizona	217	3.84	1.82	48	2.63	1.54
California	214	3.83	1.82	57	2.75	1.57
Florida	382	4.42	1.94	16	2.01	1.36
Georgia	310	4.19	1.89	7	1.65	1.25
Idaho	131	3.38	1.72	55	2.72	1.56
Illinois	242	3.94	1.84	23	2.19	1.42
Indiana	263	4.03	1.86	18	2.07	1.38
Kentucky	274	4.07	1.87	15	1.96	1.35
Louisiana	310	4.20	1.89	14	1.93	1.34
Massachusetts	333	4.27	1.91	37	2.47	1.49
Maryland	317	4.22	1.90	27	2.27	1.44
Maine	331	4.27	1.91	39	2.51	1.50
Michigan	266	4.04	1.86	25	2.24	1.43
Minnesota	192	3.72	1.79	32	2.37	1.47
Missouri	248	3.97	1.84	23	2.19	1.42
Montana	98	3.15	1.66	47	2.62	1.53
North Carolina	319	4.23	1.90	16	2.00	1.36
North Dakota	161	3.56	1.76	37	2.47	1.50
New Hampshire	328	4.25	1.90	37	2.47	1.49
New Jersey	323	4.24	1.90	31	2.37	1.47
New York	326	4.25	1.90	32	2.37	1.47
Ohio	278	4.08	1.87	20	2.13	1.40
Pennsylvania	322	4.24	1.90	29	2.33	1.46
Tennessee	273	4.06	1.86	11	1.80	1.30
Texas	290	4.13	1.88	23	2.18	1.42
Virginia	332	4.27	1.91	25	2.24	1.43
Vermont	312	4.20	1.89	37	2.47	1.49
Washington	98	3.14	1.66	64	2.82	1.59
Wisconsin	234	3.91	1.83	26	2.25	1.43

With the bilateral cost factors established we utilize the local equilibrium conditions given by A-vW (2001) to find the taste parameters and benchmark price normalization. Assume unitary output prices at the point of export ($FOB_i^* = 1, \forall i$) such that measured income is identical to the quantity endowment (e_i^0). The set of solution composite prices, CPI_i^*, must satisfy

$$CPI_i^* = \left[\sum_j \frac{e_j^0}{\sum_k e_k^0} \left(\frac{t_{ji}}{CPI_j^*} \right)^{1-s} \right]^{1/1-s}, \forall i .$$ (7)

The CES distribution parameter under homogeneous tastes (b_i^*) must be consistent with simultaneous product exhaustion across all regions;

$$e_i^0 = \sum_j e_j^0 \left[\frac{b_i^* t_{ij}}{CPI_j^*} \right]^{1-s} .$$ (8)

The benchmark parameters, CPI_i^* and b_i^*, are implicit functions of the initial income, trade costs, and s. These are solved numerically to reveal the benchmark equilibrium. The full numeric equilibrium consistent with A-vW's estimated equation is presented in Table 2, given an elasticity of substitution of five. This equilibrium also satisfies the general equilibrium conditions (1) through (4).

Table 2

<table>
<tr><td colspan="5" align="center">Benchmark Equilibrium at σ=5</td></tr>
<tr>
<td></td>
<td>Income Level
(Y; C$billions)</td>
<td>Consumer Price
Index (CPI)</td>
<td>Price of Output
(FOB)</td>
<td>Utility
(U)</td>
</tr>
<tr><td></td><td>(1)</td><td>(2)</td><td>(3)</td><td>(4)</td></tr>
<tr><td>Province</td><td></td><td></td><td></td><td></td></tr>
<tr><td>Alberta</td><td>66.07</td><td>1.98</td><td>1.00</td><td>1.00</td></tr>
<tr><td>British Columbia</td><td>75.04</td><td>1.96</td><td>1.00</td><td>1.00</td></tr>
<tr><td>Manitoba</td><td>22.44</td><td>1.96</td><td>1.00</td><td>1.00</td></tr>
<tr><td>New Brunswick</td><td>13.03</td><td>1.92</td><td>1.00</td><td>1.00</td></tr>
<tr><td>Newfoundland</td><td>8.49</td><td>2.19</td><td>1.00</td><td>1.00</td></tr>
<tr><td>Nova Scotia</td><td>16.14</td><td>1.98</td><td>1.00</td><td>1.00</td></tr>
<tr><td>Ontario</td><td>274.83</td><td>1.72</td><td>1.00</td><td>1.00</td></tr>
<tr><td>Prince Edward Island</td><td>1.92</td><td>1.98</td><td>1.00</td><td>1.00</td></tr>
<tr><td>Quebec</td><td>148.51</td><td>1.77</td><td>1.00</td><td>1.00</td></tr>
<tr><td>Saskatchewan</td><td>19.82</td><td>2.01</td><td>1.00</td><td>1.00</td></tr>
<tr><td>State</td><td></td><td></td><td></td><td></td></tr>
<tr><td>Alabama</td><td>80.35</td><td>1.51</td><td>1.00</td><td>1.00</td></tr>
<tr><td>Arizona</td><td>77.40</td><td>1.60</td><td>1.00</td><td>1.00</td></tr>
<tr><td>California</td><td>874.42</td><td>1.60</td><td>1.00</td><td>1.00</td></tr>
<tr><td>Florida</td><td>286.26</td><td>1.57</td><td>1.00</td><td>1.00</td></tr>
<tr><td>Georgia</td><td>157.97</td><td>1.50</td><td>1.00</td><td>1.00</td></tr>
<tr><td>Idaho</td><td>19.59</td><td>1.70</td><td>1.00</td><td>1.00</td></tr>
<tr><td>Illinois</td><td>308.82</td><td>1.45</td><td>1.00</td><td>1.00</td></tr>
<tr><td>Indiana</td><td>124.55</td><td>1.45</td><td>1.00</td><td>1.00</td></tr>
<tr><td>Kentucky</td><td>76.45</td><td>1.46</td><td>1.00</td><td>1.00</td></tr>
<tr><td>Louisiana</td><td>99.89</td><td>1.56</td><td>1.00</td><td>1.00</td></tr>
<tr><td>Massachusetts</td><td>187.49</td><td>1.47</td><td>1.00</td><td>1.00</td></tr>
<tr><td>Maryland</td><td>128.54</td><td>1.38</td><td>1.00</td><td>1.00</td></tr>
<tr><td>Maine</td><td>27.04</td><td>1.54</td><td>1.00</td><td>1.00</td></tr>
<tr><td>Michigan</td><td>218.93</td><td>1.45</td><td>1.00</td><td>1.00</td></tr>
<tr><td>Minnesota</td><td>113.08</td><td>1.56</td><td>1.00</td><td>1.00</td></tr>
<tr><td>Missouri</td><td>120.74</td><td>1.56</td><td>1.00</td><td>1.00</td></tr>
<tr><td>Montana</td><td>15.04</td><td>1.70</td><td>1.00</td><td>1.00</td></tr>
<tr><td>North Carolina</td><td>163.06</td><td>1.49</td><td>1.00</td><td>1.00</td></tr>
<tr><td>North Dakota</td><td>12.43</td><td>1.64</td><td>1.00</td><td>1.00</td></tr>
<tr><td>New Hampshire</td><td>28.46</td><td>1.46</td><td>1.00</td><td>1.00</td></tr>
<tr><td>New Jersey</td><td>244.22</td><td>1.16</td><td>1.00</td><td>1.00</td></tr>
<tr><td>New York</td><td>561.05</td><td>1.39</td><td>1.00</td><td>1.00</td></tr>
<tr><td>Ohio</td><td>256.62</td><td>1.45</td><td>1.00</td><td>1.00</td></tr>
<tr><td>Pennsylvania</td><td>277.35</td><td>1.41</td><td>1.00</td><td>1.00</td></tr>
<tr><td>Tennessee</td><td>108.60</td><td>1.53</td><td>1.00</td><td>1.00</td></tr>
<tr><td>Texas</td><td>423.31</td><td>1.59</td><td>1.00</td><td>1.00</td></tr>
<tr><td>Virginia</td><td>167.38</td><td>1.47</td><td>1.00</td><td>1.00</td></tr>
<tr><td>Vermont</td><td>13.17</td><td>1.52</td><td>1.00</td><td>1.00</td></tr>
<tr><td>Washington</td><td>122.81</td><td>1.71</td><td>1.00</td><td>1.00</td></tr>
<tr><td>Wisconsin</td><td>111.78</td><td>1.46</td><td>1.00</td><td>1.00</td></tr>
</table>

Data

Finally, calibrating the model requires very little external data, only regional output and bilateral distances. We use 1989 output data from Statistics Canada's Provincial Economic Accounts and the Bureau of Economic Analysis. Gross State Products were converted into Canadian dollars using the average market exchange rate for 1989. Distances are measured as the great circle distance between the principal cities of each state. Sub-regional distances are calculated as the population weighted mean distance between the three largest cities in each region. We do not require any trade data, as the model maps output directly into trade flows, given the trade cost factors reported above.

IV. Model Implications and Results of Removing Border Costs

A key element missing from the structural gravity literature is a legitimate validation of the theory. By design the theory is adept at predicting bilateral trade flows including border resistance. To make inferences beyond trade flows, however, the model should be consistent with other observable facts. The model is shown here to grossly overstate the resource cost of transport, and exaggerate the price gap generated by the Canadian-US border. These failures lead to an identified bias in welfare inferences. The estimated border costs may be too large, and welfare effects from the border are probably greatly exaggerated.

Table 3 summarizes some of the troubling characteristics of the equilibrium. We offer a wide range of substitution elasticities to strengthen our argument. Again we suggest that for this type of aggregate model the true value has a maximum of about five and is likely to be in the range of three to four given the direct transport elasticity measured by Hummels (1999). We concentrate on an elasticity of five because this is the value that A-vW assumed in their analyses [(2001) and (2000)]. Adopting higher elasticities improves some aspects of the model but considerably weakens the case that borders are important to anything other than trade flows (price differences and welfare effects are minimal at very high elasticities).

Table 3

Summary Measures for Canada and the US: Implied Price Indices, Disbursement of GDP, and Welfare Effect of Border				
	Elasticity of Substitution (σ)			
	2	5	10	40
Canada				
Weighted Average Benchmark CPI	11.3	1.8	1.3	1.1
Cost of Living Wedge Relative to the US	120%	22%	9%	2%
Tariff Equivalent of Border Barriers (US-CA)	421%	51%	20%	4%
GDP Devoted to Transport Cost (iceberg melt)	69%	44%	25%	7%
GDP Devoted to Border Cost	24%	10%	5%	1%
GDP Consumed	6%	46%	70%	92%
Welfare Cost of Border	184.7%	34.3%	14.5%	3.3%
US				
Weighted Average Benchmark CPI	5.1	1.5	1.2	1.0
Cost of Living Wedge Relative to the US	0%	0%	0%	0%
Tariff Equivalent of Border Barriers (CA-US)	421%	51%	20%	4%
GDP Devoted to Transport Cost (iceberg melt)	86%	48%	26%	7%
GDP Devoted to Border Cost	3%	1%	1%	0%
GDP Consumed	11%	51%	73%	93%
Welfare Cost of Border	10.2%	1.9%	0.8%	0.2%

The most troubling aspect of the model concerns the quantity of output that melts. This represents the real resource cost of transport. Summing across all provinces a full 54% of Canadian GDP is devoted to transport (using an elasticity of five). This includes 44% melt in transport and 10% melt at the border. In the US the model indicates that 49% of output melts, with 1% devoted to border costs. In fact, the transport sector has averaged about 3% of GDP in the US over recent decades (see the US Department of Commerce, Bureau of Economic Analysis, Industry Accounts Data, GDP by industry). Even the extreme elasticity of 40 indicates that the model over allocates output to transport by a factor of at least two.

Under the assumptions of the structural gravity model, A-vW's qualitative point that Canada, because it is smaller and has fewer varieties/regions, faces higher border costs relative to the US is supported, but the inferred quantitative magnitude of these additional costs must be

brought into question. Our data indicate a tariff equivalent rate of 51% on Canadian-US trade.[9]

Similar to transport costs, the border costs in the benchmark equilibrium are completely

determined by the interaction of distance, initial income levels, the substitution rate, and the

estimated coefficients. There is no actual attempt to measure the barriers. A-vW claim to

identify the total barrier – including formal (i.e., tariffs) and non-formal trade barriers – by

inferring the border cost from their regression. This is a tenuous method given the inability of the

model to reconcile richer more detailed data.

Table 3 also identifies a significant gap between consumer and producer prices. The cost

of living index represents the price of the composite consumption commodity. The composite

units are equivalent to those of the produced goods (which trades at a price of one in the

benchmark). This is the implicit normalization imposed by the model's symmetric solution (see

A-vW [2001] p.8). The indication is that Canadian consumer prices are 80% higher than

Canadian producer prices with a substitution elasticity of 5. It is important to keep in mind that in

this model the difference is not a markup paid to service sectors that offer products to consumers,

it is represented as a pure loss of product through transport. Because U.S. states are more

geographically central, and because a smaller share of U.S. consumption melts at the border, the

US consumer price is only 1.5 times producer prices.

The difference between the consumer price index in Canada relative to the US is further

evidence against the model structure and is at odds with the inferred border barrier. At an

elasticity of five the model indicates that it would be 22% more expensive to live in Canada

relative to the US. This is refuted by direct evidence on the cost of living. The US State

Department publishes relative cost of living statistics for cities around the world.[10] These data

place the cost of living difference in 2000, relative to Washington, DC at 9% in Vancouver, but –

---

[9] A-vW report a Canadian-US tariff equivalent rate of 49% (Table 1, A-vW [2000]). The slight difference
between our calculated rate and theirs is probably due to different distance sources or rounding of estimated
coefficients.

[10] http://www.state.gov/www/perdiems/quarterly_reports/jan00_table1.html.

3% for Ottawa. The measured similarity in the cost of living between the US and Canada directly refute the argument that barriers on the border are large.[11] Cross-border similarity in price levels also explains the absence of substantial migratory pressure, an outcome implicit in the large price differences the model predicts.

It is interesting to note that the State Department data shows little evidence for the broad claim in the literature that small countries have higher cost of living indexes because of larger border effects. For example, Japan and Germany have indexes far above the US', but many developing countries have low indexes relative to the US. It is also interesting to think about the models' implications for migration. In a model with migration and a pattern of smaller countries having higher real living costs given some border barriers, there is impetus to move to larger countries. As smaller countries lose population they become smaller, reinforcing the impetus to move. This begs the question: why has the world not evolved to one large country?

The final, and most important, results presented in Table 3 are the welfare changes induced by removing the border barriers. Using the general equilibrium system allows us to directly measure welfare effects of changes in the trade costs. This is preferable to the decomposition approach taken by A-vW (2000), because they use offline calculations to include rent-generating distortions. These are inherently inconsistent with their estimated equation (and the underlying equilibrium system they use to derive it). We find that removing the Canadian-US border increases Canadian welfare by 34.3% in the general equilibrium when the substitution elasticity is set to five. This is a spectacular result considering that researchers who actually modeled the trade and production equilibrium in Canada and the US found much smaller welfare

---

[11] Engel and Rogers (1996) document border-induced changes in the *variability* of consumer prices. Our point here is that the *level* of prices is not particularly different across borders, and that this fact contradicts the model prediction.

benefits from NAFTA (Brown, et al. [1992] for example report a 0.7% increase in Canadian welfare).[12]

Explaining the welfare increase in the gravity model is quite easy. In the benchmark 10% of income was devoted to border costs. Removing these costs initially increases income by 10%. As consumers adjust to their optimal bundle, welfare goes up by more than 10%, and this substitution effect is reinforced by an increase in demand for Canadian output by US consumers. The percentage changes in all the variables of the equilibrium system when we remove the border costs are presented in Table 4 (we arbitrarily selected the output price of a US state, Alabama, as the numeraire).

---

[12] The differences between the structural gravity approach and applied general equilibrium results should not be surprising. Computational general equilibrium models are based on limited data and identifying assumptions – precisely what the structural gravity model is based on. In the structural gravity model homogeneous preferences are an identifying assumption. This is costly because it rules out richer theories of production location effects, etc. In simulation models heterogeneous tastes in combination with measured policy instruments replicate flows consistent with the border effect, but the model is still open to any general trade structure. Any non-measured trade barriers at the border, by construction, have no welfare implications, because agents are reacting to the instruments they face – and only those instruments.

Table 4.

	Scenario Equilibrium (no border barriers) at σ=5			
	Income Level (percent change)	**Consumer Price Index (percent change)**	**Price of Output (percent change)**	**Utility (percent change)**
	(1)	(2)	(3)	(4)
Province				
Alberta	12.44%	-14.78%	12.44%	31.95%
British Columbia	13.14%	-15.44%	13.14%	33.80%
Manitoba	14.43%	-16.62%	14.43%	37.24%
New Brunswick	13.61%	-15.87%	13.61%	35.05%
Newfoundland	15.31%	-17.42%	15.31%	39.64%
Nova Scotia	14.35%	-16.55%	14.35%	37.03%
Ontario	13.54%	-15.81%	13.54%	34.86%
Prince Edward Island	13.59%	-15.86%	13.59%	35.01%
Quebec	12.80%	-15.12%	12.80%	32.90%
Saskatchewan	14.31%	-16.52%	14.31%	36.93%
State				
Alabama	0.00%	-1.33%	0.00%	1.34%
Arizona	0.01%	-1.33%	0.01%	1.36%
California	-0.04%	-1.28%	-0.04%	1.26%
Florida	-0.10%	-1.21%	-0.10%	1.13%
Georgia	0.04%	-1.37%	0.04%	1.43%
Idaho	0.77%	-2.27%	0.77%	3.11%
Illinois	0.21%	-1.59%	0.21%	1.83%
Indiana	0.18%	-1.54%	0.18%	1.75%
Kentucky	0.12%	-1.48%	0.12%	1.62%
Louisiana	0.00%	-1.32%	0.00%	1.34%
Massachusetts	0.68%	-2.16%	0.68%	2.90%
Maryland	0.19%	-1.56%	0.19%	1.78%
Maine	1.21%	-2.80%	1.21%	4.12%
Michigan	0.81%	-2.32%	0.81%	3.21%
Minnesota	0.48%	-1.91%	0.48%	2.43%
Missouri	0.38%	-1.79%	0.38%	2.21%
Montana	1.07%	-2.63%	1.07%	3.80%
North Carolina	0.11%	-1.46%	0.11%	1.59%
North Dakota	0.89%	-2.42%	0.89%	3.39%
New Hampshire	0.63%	-2.10%	0.63%	2.80%
New Jersey	-0.46%	-0.75%	-0.46%	0.29%
New York	0.38%	-1.79%	0.38%	2.21%
Ohio	0.93%	-2.46%	0.93%	3.47%
Pennsylvania	0.34%	-1.75%	0.34%	2.12%
Tennessee	0.09%	-1.44%	0.09%	1.55%
Texas	-0.01%	-1.32%	-0.01%	1.33%
Virginia	0.22%	-1.59%	0.22%	1.84%
Vermont	2.61%	-4.45%	2.61%	7.38%
Washington	1.98%	-3.71%	1.98%	5.91%
Wisconsin	0.29%	-1.68%	0.29%	2.00%

Table 5 shows the significant changes in gross trade flows. We measure the flow in value ($FOB_i X_{ij}$). All goods are traded so the sum of the exports equals income. Trade across the Canadian-US border is predicted to increase by about 200% with a substitution elasticity of 5. Income increases in Canada because gross demand rises for its products. The value of trade within Canada falls by 65%. The model predicts significant swings in trade flows because it requires substantial price differences to fit the observed trade data.

Table 5.

Value of Trade Flows at $\sigma=5$			
	To Canada	To the US	Total (GDP)
Exports from Canada			
Benchmark	451	195	646
Counterfactual (no border cost)	158	575	732
Change	-65.1%	194.7%	13.3%
Exports from US			
Benchmark	195	5212	5407
Counterfactual (no border cost)	575	4846	5420
Change	194.7%	-7.0%	0.25%

The structural gravity framework places all of the border resistance observed in the trade data on the ad hoc tariff equivalent border cost. There is no structural explanation of the non-formal trade barrier, only anecdotal references to differences in cultures, customs, and regulations. A careful approach that models the underlying cause of border resistance will likely lead to very different conclusions. For example, cultural differences might legitimately be modeled as heterogeneous tastes. In this case the border resistance results from consumer's preference for the home good (the "voting with your feet" literature supports spatial agglomeration of like individuals around places where their preferred goods are cheapest, Tiebout [1956]). If people simply prefer the home good, observed trade resistance at a border with no formal tariffs does not necessarily indicate any welfare effect of the border. Obviously, non-traded goods, intermediate trade, and standard endogenous plant location models could have

similar implications. These richer theories are more attractive than the structural gravity model, which identifies border resistance without explaining it.

V. Lessons and A Constructive Look at Future Research

We view the results presented above as a precautionary tale about the shortcomings of overly simple models. While parsimony is useful for expositional purposes, its value in policy analysis and welfare estimation can be limited. Goods differ in ways that matter for understanding trade patterns. Input requirements, tradability, and the geographic distribution of demand differ across goods, and these distinctions have important implications for welfare measurement. Models predicting a gravity-like trade pattern in aggregate trade are best viewed as metaphors, not precise analytical tools.

Structural gravity models fail in three important ways. First, they ignore the role of intermediates in trade. This failure is especially important in North American interregional trade patterns, where trade in intermediates is pervasive. Low trade costs induce co-location of firms with up- and down-stream suppliers, a theoretical point raised by Krugman and Venables (1996). Second, the models misrepresent the nature and costs of transportation by failing to model the transportation sector directly. The iceberg formulation treats transportation costs as a pure evaporation of product. Transport is better viewed as a real economic activity in which factors are paid their opportunity costs in other sectors. Third, the models fail to account for nontradeables, or in cross- commodity variation of tradability across geographic space.

Where does our analysis leave the gravity model? We believe that the empirical gravity relationship has a limited, but useful analytical role. It is most valuable for making inferences when a trade policy regime shift is expected. Analyses of likely trade patterns in post-communist Eastern Europe (Hamilton and Winters, 1992) and post-sanctions Cuba (U.S. International Trade Commission, 2001) benefit from the gravity model's predictive power. We note that these are

carefully circumscribed experiments that rely on the empirical gravity relationship because historical trade data are not applicable to the relevant counterfactual.

Future empirical work on the geographic pattern of trade should emphasize commodity disaggregation of trade flows. The commodity composition of trade varies over geographic space. To the extent that there are credible micro-foundations of the aggregate gravity model, they are likely to lie in regularities of cross-commodity aggregation. Haveman and Hummels (1999) offer useful insights here, but more work is needed. Hummels (1999) offers helpful guidance about the appropriate econometric methods for use with disaggregated data.

The economic geography literature has been overly reliant on trade flow data for its inferences about the value of economic linkages. The welfare losses of missing trade are best estimated through spatial deviations in prices. Imputed prices, like those in A-vW, are very poor substitutes for actual data. We recognize that in many cases, suitable price data are difficult to find. We expect that researchers uncovering such data would provide the literature with a number of valuable insights.

Finally, we note that calibration exercises are a useful approach to better understanding these issues. Dawkins, et al. (2001) make the point: "estimation is calibration and calibration is estimation". The two empirical tools do rely on different identification strategies, and these should be judged as more or less appropriate for a given exercise. In the case of the aggregate gravity model, the implicit identifying assumptions in structural econometric work (i.e. homogenous consumer preferences and no production location effects associated with intermediate goods trade) lead the analyst to conclude that there must be large border frictions that substantially distort prices. A calibration exercise, even one using the Anderson framework, would not make the same implausible inferences. Calibration exercises would emphasize the size of "home bias," rather than high border costs. This alternate understanding of the border effect produces far lower estimates of border-related welfare costs and a more plausible representation of geographic price variation. A more sophisticated calibration exercise, such as a Computable

General Equilibrium (CGE) model that included input-output tables, would be even better able to represent the real world.   It is unfortunate that calibrated CGE models can be so easily dismissed as "black boxes." Willful ignorance of the CGE literature is not likely to be a productive research strategy for those who would credibly estimate the welfare effects of trade interruptions.

V. Conclusion

Despite its good empirical fit, the micro-foundations of the gravity model remain a mystery. The existing class of models that "explain" gravity rely too heavily on price differentials and consumer responses to those differentials. The models explain trade well, of course, but that is no surprise. The robust empirical relationship they predict was known to explain the trade pattern well before the models were written down. It is only through assessing the models' general equilibrium predictions that their weaknesses become apparent.

The models' weaknesses lie in their combination of four doubtful assumptions: 1) all output is tradable 2) all goods are produced in a unique location, 3) there is no intermediate goods trade, and 4) trade frictions take the ad valorem "iceberg" form.   Strategic simplification is necessary in modeling, of course. But the combined impact of these assumptions is sufficient to render these models unsuitable for welfare estimation or policy analysis.

To date, the weaknesses of the gravity model have not been highlighted because the model has not been simulated in full general equilibrium. Our simulations, which are based on the econometric work of the models' greatest champions, show that the theoretical models lean far too heavily on transport costs and their subsequent effect on consumer behavior.   A careful reading of earlier work reveals that the price differences implied by structural models are far too large. Our work shows why. Under plausible parameterizations of a standard gravity model, we find that transport costs must consume over 45% of output if the model is to explain interregional trade patterns in the U.S. and Canada.

Our results cast severe doubt on the value of structural gravity models for measuring welfare. We believe a more nuanced approach is required. Empirical work should focus on estimating relationships at the detailed commodity level. Studies that used price data explicitly, rather than constructed proxy variables, would be welcomed. Simulation work should include more thorough modeling of the transport sector. Careful consideration of input-output relationships, and their effect on variables of interest to the economic geography literature, is also in order.

At a minimum, we hope we have discredited the use of structural gravity models for welfare measurement and policy analysis. Pressed to "take the theory seriously," we have done so. We find the theory wanting in several important respects, especially those that are most important for measuring welfare. We urge care in advertising the quantitative welfare implications of these stylized models, especially in policy arenas.

Bibliography

Anderson, James E. 1979. "A Theoretical Foundation for the Gravity Equation" *American Economic Review* 69(March): 106-116.

Anderson, James and Eric van Wincoop 2001. "Gravity with Gravitas: A Solution to the Border Puzzle," NBER Working Paper 8079, (January).

Anderson, James and Eric van Wincoop 2000. "Borders, Trade and Welfare," paper prepared for the Brookings Trade Forum 2001 on *Globalization: Issues and Implications*, (May 10-11, 2001).

Anderson, Michael A., and Stephen L. S. Smith 1999a. "Canadian Provinces in World Trade: Engagement and Detachment" *Canadian Journal of Economics* 32 (February): 22-38.

Anderson, Michael A., and Stephen L. S. Smith 1999b. "Do National Borders Really Matter? Canada-U.S. Regional Trade Reconsidered" *Review of International Economics* 7 (May): 219-227.

Bergstrand, Jeffrey H. 1985. "The Gravity Equation in International Trade: Some Microeconomic Foundations and Empirical Evidence." *Review of Economics and Statistics* 67 (August): 474-481.

Brown, Drusilla K., Alan V. Deardorff, and Robert M. Stern, 1992 "A North American Free Trade Agreement: Analytical Issues and a Computational Assessment," *World Economy* 15, pp. 11-29.

Dawkins, Christina, T.N. Srinivasan, and John Whalley, 2001, "Calibration," *Handbook of Econometrics*, Volume 5. J.J. Heckman and E. Leamer, eds, Elsevier Science, pp 3655-3701.

Deardorff, Alan V. 1998. "Determinants of Bilateral Trade: Does Gravity Work in a Neoclassical World?" in *The Regionalization of the World Economy*, Jeffrey A. Frankel ed. Cambridge, MA: National Bureau of Economic Research. 7-22.

Engel, Charles and John Rogers. 1996. "How Wide is the Border?" *American Economic Review* 86 (December) pp. 1112-1125.

Evans, Carolyn 1999. *Do National Borders Matter?* Doctoral Dissertation, Harvard University.

Evans, Carolyn 2001, "Big border effects, small border costs?" New York Federal Reserve, *mimeo.*

Gallaway, Michael, Christine McDaniel ands Sandra Rivera 2000. "Industry-Level Estimates of U.S. Armington Elasticities." U.S. International Trade Commission Working Paper, #2000-09-A, (September)

Hamilton, Carl and L. Alan Winters 1992. "Opening up to International Trade in Eastern Europe," *Economic Policy*, (April), pp. 78-116.

Haveman, John and David Hummels 1999. "Alternative Hypotheses and the Volume of Trade: Evidence on the Extent of Specialization" University of Chicago *mimeo* (January).

Helliwell, John F. 1996. "Do National Borders Matter for Quebec's Trade?" *Canadian Journal of Economics* 29 (August): 507-522.

Hillberry, Russell 1998. "Regional Trade and the 'Medicine Line': The National Border Effect in U.S. Commodity Flow Data," *Journal of Borderlands Studies*, (Fall), pp. 1-17.

Hillberry, Russell 2001. "Aggregation Bias, Compositional Change, and the Border Effect," U.S. International Trade Commission Working Paper #2001-04-B, (April).

Hillberry, Russell and David Hummels 2000. "Explaining Home Bias in Consumption: Production Location, Commodity Composition and Magnification," Purdue University *mimeo* (December).

Hummels, David 1999. "Toward a Geography of Trade Costs," University of Chicago *mimeo* (January).

Krugman, Paul 1980 "Scale Economies, Product Differentiation, and the Pattern of Trade," *American Economic Review,* 70 (December): 950-959.

Krugman Paul and Anthony Venables 1996. "Integration, Specialization, and Adjustment." *European Economic Review* 40 (April): 959-967.

McCallum, John T. 1995. "National Borders Matter: Canada-U.S. Regional Trade Patterns*."* *American Economic Review* 85 (June): 615-623.

Obstfeld, Maurice and Kenneth Rogoff 2000. "The Six Major Puzzles in Modern Macroeconomics: Is There a Common Cause?" NBER Working Paper 7777, (July).

Rutherford, T.F. (1995) "Extensions of GAMS for Complementarity Problems Arising in Applied Economics," *Journal of Economic Dynamics and Control*, 1299-1324.

Tiebout, C. M. (1956) "A Pure Theory of Local Expenditures," Journal of Political Economy, 64, October: 416-424.

U.S. International Trade Commission 2001. *The Economic Impact of Sanctions With Respect to Cuba,* USITC publication 3398. (February).

Wei, Shang-Jin. 1996. "Intra-national Versus International Trade: How Stubborn Are Nations in Global Integration?" NBER Working Paper 5531. (April).